Efficient Deterministic Conditional Sampling of Multivariate Gaussian Densities

Daniel Frisch and **Uwe D. Hanebeck**

***Abstract*— We propose a fast method for deterministic multivariate Gaussian sampling. In many application scenarios, the commonly used stochastic Gaussian sampling could simply be replaced by our method – yielding comparable results with a much smaller number of samples. Conformity between the reference Gaussian density function and the distribution of samples is established by minimizing a distance measure between Gaussian density and Dirac mixture density. A modified Cramér-von Mises distance of the Localized Cumulative Distributions (LCDs) of the two densities is employed that allows a direct comparison between continuous and discrete densities in higher dimensions. Because numerical minimization of this distance measure is not feasible under real time constraints, we propose to build a library that maintains sample locations from the standard normal distribution as a template for each number of samples in each dimension. During run time, the requested sample set is re-scaled according to the eigenvalues of the covariance matrix, rotated according to the eigenvectors, and translated according to the mean vector, thus adequately representing arbitrary multivariate normal distributions.**

I. Introduction

Context: Gaussian sampling is a core component of nonlinear Gaussian filtering [1], [2], [3], [4], [5], [6] and Gaussian sum filtering [7], [8], furthermore in particle filters for proposal density sampling [9], [10] or smoothing [11], in sequential Monte Carlo methods [12], in control scenarios [13], [14], and many kinds of simulations [15], [16], [17], [18]. For mixed Euclidean and non-Euclidean manifolds like the Special Euclidean groups, normally distributed samples are necessary for the Euclidean part [19], [20]. Even in purely non-Euclidean Riemannian manifolds, Gaussian samples can be used in the Euclidean tangent space and subsequently projected on the manifold [21], [22], [23].

In lots of cases, random Gaussian samples can be replaced by much fewer deterministic samples. This reduces computational complexity, especially if the nonlinear system equations involve expensive computations like numerical integration.

Considered Problem: In this work we consider the problem of deterministic Gaussian sampling. We present a method that 1) is very fast and 2) provides a sampling quality that is suboptimal but sometimes better than existing methods.

State-of-the-art: The simplest Gaussian sampling scheme is employed by the Unscented Kalman Filter (UKF) with just two samples on each principal axis [24], [25]. Estimation quality can be improved by placing more than two samples on the principal axes [6]. Sample locations are thereby obtained by minimizing the L_2 distance between the respective scalar *cumulative* distribution functions of Gaussian and Dirac mixture (DM). Even better is to distribute the samples freely in space, wherever probability mass of the desired Gaussian function is located. Cumulative distributions of multivariate random variables are however ambiguous in their definition. One possibility to overcome this is to take all possible projections into account [26].

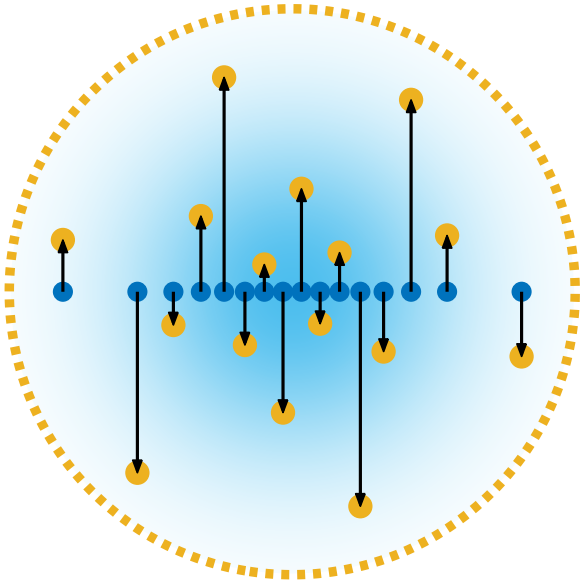


Fig. 1: Visualization of the proposed deterministic sampling scheme for a bivariate standard normal density (blue shade), starting from the univariate case. First, a one-dimensional standard normal distribution is sampled (1) on the x-axis (blue points). Second, a two-dimensional standard normal density (yellow points) is sampled (2) by adding y-coordinates (black arrows) to the existing samples. Subsequent re-scaling (4) against the arrows to obtain arbitrary Gaussians (not shown here) retains much of the uniformity of sample distribution.

In this paper, we instead exploit the properties of the Localized Cumulative Distribution (LCD) as a unique continuous representation of multivariate continuous or discrete densities [27]. Using the LCD for the specific case of approximating multivariate Gaussian densities with DM densities gives very good results [28], [29]. However, it is computationally expensive, which prohibits online application, e.g., in Gaussian filtering. The Smart Sampling Kalman Filter (S^2KF) circumvents this problem by using an offline generated library of standard normal distributions [1], [30], [31]. An

Daniel Frisch and Uwe D. Hanebeck are with the Intelligent Sensor-Actuator-Systems Laboratory (ISAS), Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany {daniel.frisch, uwe.hanebeck}@ieee.org.

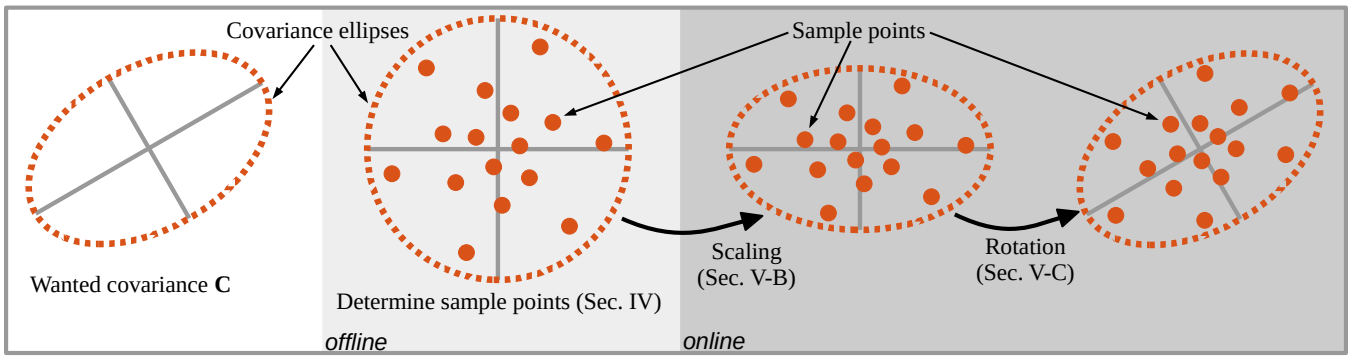


Fig. 2: For obtaining deterministic conditional samples with arbitrary covariance $\mathbf{C}$, at first the sample points for the standard normal distribution have to be determined offline. Performing the scaling and rotation operations online completes the approximation. A similar figure for Gaussian sampling only on principal axes appeared in [6].

open source implementation of the S^{2}KF and its library-based Gaussian sampling is available [32].

Open Problems: Problem remains with the sampling approach used in the S^{2}KF that the online transformation of the standard normal samples often leads to unwanted sample aggregation such that the distribution of points is not very uniform. Furthermore, the severity of this behaviour depends not only on the eigenvalues of the covariance matrix, what can be expected, but also on its eigenvectors, i.e., the rotational part. This behaviour is undesired.

II. Key Idea

We propose a standard normally distributed set of samples in the N-dimensional Euclidean space that is designed to be re-scaled only along the N coordinate axes, see Fig. 1. The resulting "elliptic" (but still axis aligned) point cloud can subsequently be rotated and therefore represent correlated normal distributions likewise, see Fig. 2.

Said template is generated by deterministic conditional sampling, i.e., samples are only moved on lines in parallel to the coordinate axes. This has several advantages. First, the wanted "uniformity" of samples is mostly retained when the point cloud is subsequently re-scaled (along the coordinate axes). As a welcome side effect, computational costs to compute the template are reduced. Coordinates of each dimension can be optimized separately, as coordinates of lower dimensions do not change anymore when higher dimensions are added. This makes the numerical optimization easier and decreases memory requirements for sample libraries.

III. Overview

Our method provides real-time capable deterministic Gaussian sampling. Required inputs are

- I1 the number L of wanted samples,
- I2 eigenvalues of the covariance matrix,
- I3 eigenvectors of the covariance matrix,
- I4 and the mean vector.

Based on these, and a previously generated offline sample library, we compute a deterministic set of particles that well approximates the thereby specified normal distribution.

We will now in Sec. IV focus on the generation of the offline library of standard normal samples. After that, in Sec. V we describe how during runtime deterministic samples for any (non-standard) Gaussian density can be obtained quickly. Sec. VI details the distance measure we use to compare normal distributions to DMs. We also present visual examples in an evaluation, Sec. VII.

IV. Offline Conditional Sampling

Consider a *standard* multivariate Gaussian density ${}^{\mathrm{G}}f^{\mathrm{s}}_N$ with $\underline{x}^{(1:N)} = \begin{bmatrix} x^{(1)} & x^{(2)} & \dots & x^{(N)} \end{bmatrix}^\top \in \mathbb{R}^N$

$$
\begin{aligned}
{}^{\mathrm{G}}f^{\mathrm{s}}_N\left(\underline{x}^{(1:N)}\right) &= \mathcal{N}\left(\underline{x}^{(1:N)} - \underline{0}\,,\ \mathbf{I}\right) \\
&= \frac{1}{\sqrt{(2\pi)^N}} \exp\left\{-\frac{1}{2}\left\|\underline{x}^{(1:N)}\right\|_2^2\right\} \ .
\end{aligned}
$$

Deterministic samples from this density are to be stored in an offline library. We will now describe the individual steps we propose to draw these samples.

A. One-Dimensional Sampling

According to our conditional sampling scheme, a univariate problem is considered initially. Note that the approximation quality along this "first" dimension will be superior to the following ones; therefore it will be assigned the biggest eigenvalue later during covariance adaptation in Sec. V-B.

We approximate the scalar Gaussian density ${}^{\mathrm{G}}f^{\mathrm{s}}_1(x)$

$$
{}^{\mathrm{G}}f^{\mathrm{s}}_1(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\,x^2\right\}
$$

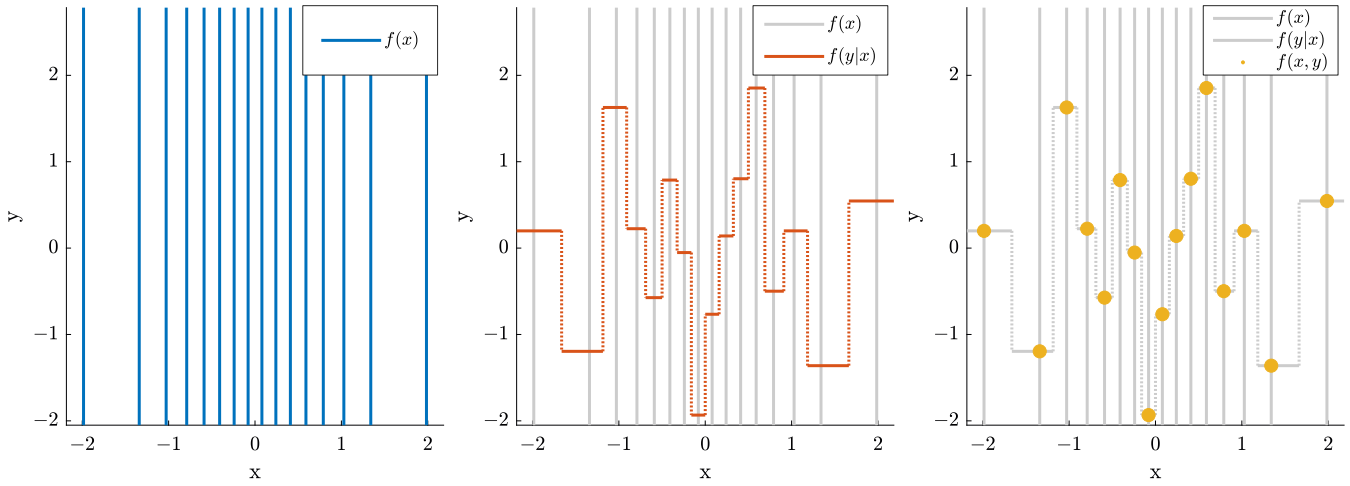
with an unweighted DM density ${}^{\mathrm{DM}}f^{\mathrm{s}}_1(x)$ with L samples x_i

$$
{}^{\mathrm{DM}}f^{\mathrm{s}}_1(x) = \frac{1}{L}\sum_{i=1}^{L} \delta(x - x_i) \ .
$$

The approximation is performed by finding the L sample locations x_i such that a distance measure $D_1(\cdot,\cdot)$ for density functions is minimized

$$
\{x_i\}_{i=1}^{L} = \arg\min_{x_i} \left\{D_1\left({}^{\mathrm{G}}f^{\mathrm{s}}_1(x),\ {}^{\mathrm{DM}}f^{\mathrm{s}}_1(x)\right)\right\} \ . \qquad (1)
$$

The resulting sample spacings are in accordance with the desired underlying continuous density, i.e., samples are closer



(a) Initial marginal distribution (b) Conditional distribution (one possibility) (c) Final joint distribution, product of (a) and (b)

Fig. 3: Visualization of the formal procedure determining the deterministic standard normally distributed samples: Adding y-coordinates to a given marginal distribution $f(x)$ can be written as multiplication with the conditional density $f(y|x)$ (3). Note that there are conditional distributions other than the one shown in (b) that would yield the same joint distribution (c).

in the region of high density and further apart for regions of low density, see blue points in Fig. 1.

The distance measure $D_n(\cdot,\cdot)$ will be described later. Important for now is that the distance measure serves as a nonlinear objective function for the L unknown parameters x_i. Thus, the optimal sample locations can be found using nonlinear optimization.

B. Two-Dimensional Problem

Now we consider the two-dimensional standard normal distribution ${}^{\mathrm{G}}f_2^{\mathrm{s}}(x,y)$

$${}^{\mathrm{G}}f_2^{\mathrm{s}}(x,y) = \frac{1}{2\pi}\exp\left\{-\frac{1}{2}\left(x^2+y^2\right)\right\}$$

and the two-dimensional DM density ${}^{\mathrm{DM}}f_2^{\mathrm{s}}(x,y)$ with L samples $\begin{bmatrix} x_i & y_i \end{bmatrix}^\top$

$${}^{\mathrm{DM}}f_2^{\mathrm{s}}(x,y) = \frac{1}{L}\sum_{i=1}^{L}\delta\left(\begin{bmatrix}x\\y\end{bmatrix} - \begin{bmatrix}x_i\\y_i\end{bmatrix}\right) \quad .$$

The sample x-values x_i as defined by (1) are inserted here as constants and *not* modified anymore. Only the L unknown y-vales y_i along the second dimension are computed by solving

$$\{y_i\}_{i=1}^{L} = \arg\min_{y_i}\left\{D_2\left({}^{\mathrm{G}}f_2^{\mathrm{s}}(x,y),\ {}^{\mathrm{DM}}f_2^{\mathrm{s}}(x,y)\right)\right\} \quad . \quad (2)$$

Clearly the samples $\begin{bmatrix} x_i & y_i \end{bmatrix}^\top$ move on lines (black arrows in Fig. 1) along the y-direction until the overall arrangement corresponds to a two-dimensional standard normal distribution (yellow points in Fig. 1).

Formally, the augmentation of ${}^{\mathrm{DM}}f_1^{\mathrm{s}}(x)$ with y-coordinates yielding the joint density ${}^{\mathrm{DM}}f_2^{\mathrm{s}}(x,y)$

$${}^{\mathrm{DM}}f_2^{\mathrm{s}}(x,y) = {}^{\mathrm{DM}}f_{2|1}^{\mathrm{s}}(y|x)\cdot{}^{\mathrm{DM}}f_1^{\mathrm{s}}(x) \quad (3)$$

can be seen as multiplication with a conditional density

$${}^{\mathrm{DM}}f_{2|1}^{\mathrm{s}}(y|x) = \begin{cases}\alpha_i\cdot\delta(y-y_i)\,, & x = x_i\\ \text{undef}\,, & \text{other } x\end{cases}$$

with a_i such that $\int_{\mathbb{R}}{}^{\mathrm{DM}}f_{2|1}^{\mathrm{s}}(y|x_i)\,\mathrm{d}y = 1,\ i\in[1,\ldots,L]$, see Fig. 3 for a visualization.

For representing a *non*-standard normal distribution later during runtime, the points can easily be moved along the same lines (black arrows in Fig. 1). This results in a relatively smooth point cloud without forming "lumps". For example, an uncorrelated normal distribution with $\sigma_x = 1$ and σ_y, with $0\le\sigma_y\le 1$, would be represented as

$${}^{\mathrm{DM}}f_2(x,y) = \frac{1}{L}\sum_{i=1}^{L}\delta\left(\begin{bmatrix}x\\y\end{bmatrix} - \begin{bmatrix}x_i\\\sigma_y\cdot y_i\end{bmatrix}\right) \quad . \quad (4)$$

Even in the extreme case of $\sigma_y = 0$ (positive semidefinite covariance matrix), the samples would not "clump" but rather optimally arrange side by side as defined by (1), see blue points in Fig. 1. The samples are thus "backward compatible" to lower dimensions.

C. N-Dimensional Problem

Extending this scheme to any dimension N is straightforward. It can be formulated as a recursion from dimension n to $(n+1)$. The start of the recursion has been described in Sec. IV-A ($n=0$) and in Sec. IV-B ($n=1$). The goal here is the DM approximation of the $(n+1)$-dimensional standard Gaussian density ${}^{\mathrm{G}}f_{n+1}^{\mathrm{s}}$

$${}^{\mathrm{G}}f_{n+1}^{\mathrm{s}}\left(\underline{x}^{(1:n+1)}\right) = \frac{\exp\left\{-\frac{1}{2}\left(\sum_{k=1}^{n+1}\left(x^{(k)}\right)^2\right)\right\}}{\sqrt{(2\pi)^{n+1}}}$$

with an $(n+1)$-dimensional DM density with L samples

$$^{\mathrm{DM}}f^{\mathrm{s}}_{n+1}\Big(\underline{x}^{(1:n+1)}\Big) = \frac{1}{L}\sum_{i=1}^{L}\delta\Big(\underline{x}^{(1:n+1)} - \underline{x}_i^{(1:n+1)}\Big) \quad .$$

Now suppose an n-dimensional DM approximation

$$^{\mathrm{DM}}f^{\mathrm{s}}_{n}\Big(\underline{x}^{(1:n)}\Big) = \frac{1}{L}\sum_{i=1}^{L}\delta\Big(\underline{x}^{(1:n)} - \underline{x}_i^{(1:n)}\Big)$$

of the n-dimensional marginal distribution

$$^{\mathrm{G}}f^{\mathrm{s}}_{n}\Big(\underline{x}^{(1:n)}\Big) = \int_{-\infty}^{\infty} {}^{\mathrm{G}}f^{\mathrm{s}}_{n+1}\Big(\underline{x}^{(1:n+1)}\Big)\ \mathrm{d}x^{(n+1)}$$

is already given.

Similar to Sec. IV-A and Sec. IV-B, there are L unknown scalar coordinates $x_i^{(n+1)}$ defined by

$$\Big\{x_i^{(n+1)}\Big\}_{i=1}^{L} = \underset{x_i^{(n+1)}}{\arg\min} \ \lrcorner \quad \Big\{D_{n+1}\Big({}^{\mathrm{G}}f^{\mathrm{s}}_{n+1}\Big(\underline{x}^{(1:n+1)}\Big),\ {}^{\mathrm{DM}}f^{\mathrm{s}}_{n+1}\Big(\underline{x}^{(1:n+1)}\Big)\Big)\Big\} \quad . \tag{5}$$

They can be obtained via nonlinear optimization.

The formal perspective of this procedure involves a certain conditional density that augments another dimension

$$^{\mathrm{DM}}f^{\mathrm{s}}_{n+1}\Big(\underline{x}^{(1:n+1)}\Big) \ \lrcorner \\ = {}^{\mathrm{DM}}f^{\mathrm{s}}_{n+1|n}\Big(\underline{x}^{(n+1)}|\underline{x}^{(1:n)}\Big)\cdot {}^{\mathrm{DM}}f^{\mathrm{s}}_{n}\Big(\underline{x}^{(1:n)}\Big) \quad .$$

Note that this conditional density $^{\mathrm{DM}}f^{\mathrm{s}}_{n+1|n}$ is not uniquely defined for DM densities: the multiplication in (5) cancels all parts where $^{\mathrm{DM}}f^{\mathrm{s}}_{n}$ has no support, see Fig. 3. The general conditional density $^{\mathrm{DM}}f^{\mathrm{s}}_{n+1|n}$ is never explicitly calculated in our method. We only determine it where $^{\mathrm{DM}}f^{\mathrm{s}}_{n}$ has support.

V. Real-Time Covariance Adaptation

This section will describe how samples for any (non-standard) normal distribution can quickly be obtained by applying a linear transformation to the standard normally distributed reference samples from Sec. IV. The process of modifying the covariance matrix of a set of samples is also called covariance adaptation. Additionally, a nonzero mean vector can be impressed.

A. Arbitrary Multivariate Normal Distribution

The continuous Gaussian density $^{\mathrm{G}}f_N$ to be approximated is given as a general multivariate normal distribution

$$\begin{aligned} &{}^{\mathrm{G}}f_N\Big(\underline{x}^{(1:N)}\Big) \\ &= \mathcal{N}\Big(\underline{x}^{(1:N)} - \underline{\mu},\ \mathbf{C}\Big) \\ &= \frac{\exp\Big\{-\frac{1}{2}\left(\underline{x}^{(1:N)} - \underline{\mu}\right)^{\top}\mathbf{C}^{-1}\left(\underline{x}^{(1:N)} - \underline{\mu}\right)\Big\}}{\sqrt{(2\pi)^N\det(\mathbf{C})}} \quad , \end{aligned}$$

with $\underline{x}^{(1:N)} \in \mathbb{R}^N$, mean vector $\underline{\mu} \in \mathbb{R}^N$, and general non-diagonal, positive semidefinite covariance matrix $\mathbf{C} \in \mathbb{R}^{N\times N}$. The covariance matrix can be decomposed according to

$$\mathbf{C} = \mathbf{R}\mathbf{\Sigma}\mathbf{R}^{\top} \tag{6}$$

into i) an orthogonal (i.e., pure rotation) matrix $\mathbf{R}$ consisting of the eigenvectors of $\mathbf{C}$, and ii) a diagonal matrix $\mathbf{\Sigma}$

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \sigma_N^2 \end{bmatrix}$$

that contains the eigenvalues, i.e., the variances σ_n^2 along the main axes that are spanned by the columns of $\mathbf{R}$. The variances are sorted from large to small, $\sigma_n^2 \geq \sigma_{n+1}^2$.

B. Anisotropic but Axis Aligned Scaling

Let the standard normally distributed DM for L samples in N dimensions be given by

$${}^{\mathrm{DM}}f^{\mathrm{s}}_{N}\left(\begin{bmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(N)} \end{bmatrix}\right) = \frac{1}{L}\sum_{i=1}^{L}\delta\left(\begin{bmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(N)} \end{bmatrix} - \begin{bmatrix} x_i^{(1)} \\ x_i^{(2)} \\ \vdots \\ x_i^{(N)} \end{bmatrix}\right)$$

from the offline library as described in Sec. IV. We re-scale this point cloud, see Fig. 2, by moving the samples back along the same lines they were moved during construction by conditional sampling. The resulting DM

$${}^{\mathrm{DM}}f^{\mathrm{u}}_{N}\left(\begin{bmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(N)} \end{bmatrix}\right) = \frac{1}{L}\sum_{i=1}^{L}\delta\left(\begin{bmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(N)} \end{bmatrix} - \begin{bmatrix} \sigma_1\, x_i^{(1)} \\ \sigma_2\, x_i^{(2)} \\ \vdots \\ \sigma_N\, x_i^{(N)} \end{bmatrix}\right)$$

now approximates the *uncorrelated* Gaussian density $^{\mathrm{G}}f^{\mathrm{u}}_{N}$

$${}^{\mathrm{G}}f^{\mathrm{u}}_{N}\Big(\underline{x}^{(1:N)}\Big) = \mathcal{N}\Big(\underline{x}^{(1:N)} - \underline{0},\ \mathbf{\Sigma}\Big) \quad .$$

If the standard deviations σ_n are not all the same, this operation is *anisotropic* re-scaling. Note that the approximation quality generally deteriorates from anisotropic re-scaling. Only the standard normally distributed point locations are actually optimized. However the proposed conditional sampling scheme pursued here keeps the loss of approximation quality at a tolerable minimum while providing an invaluable saving of computation time.

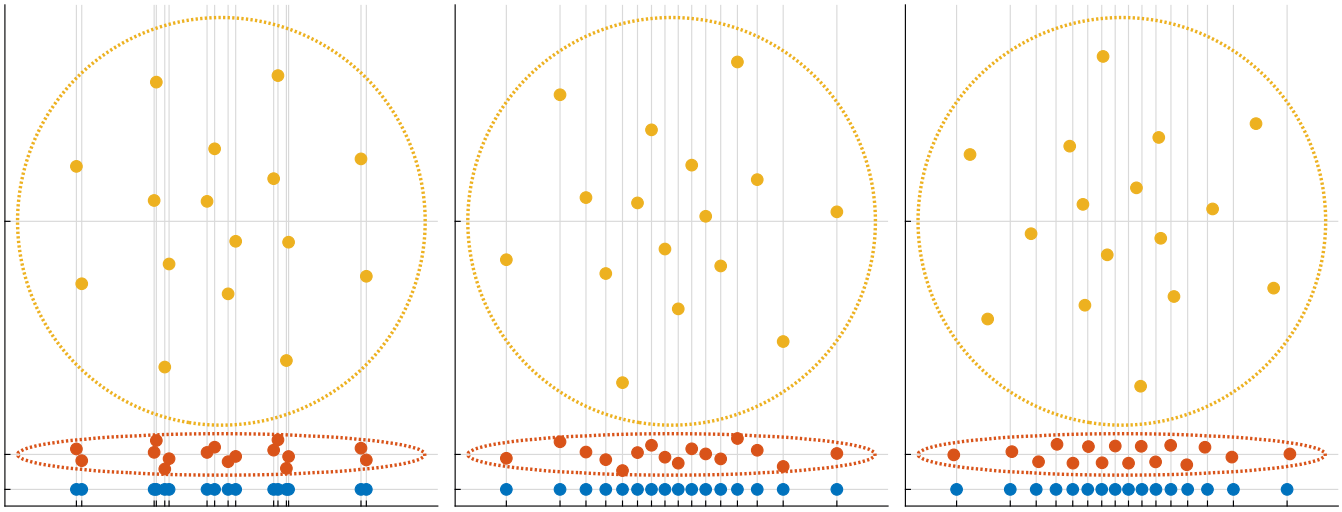
C. Rigid Transformation

Finally, the now "elliptical" but still axis aligned point cloud is *rotated*, see Fig. 2, to represent the potentially non-diagonal covariance matrix $\mathbf{C}$. In addition, it is translated by the mean vector $\underline{\mu}$. This results in $^{\mathrm{DM}}f_N$

$${}^{\mathrm{DM}}f_N\Big(\underline{x}^{(1:N)}\Big) = \frac{1}{L}\sum_{i=1}^{L}\delta\left(\underline{x}^{(1:N)} - \mathbf{R}\begin{bmatrix} \sigma_1\, x_i^{(1)} \\ \sigma_2\, x_i^{(2)} \\ \vdots \\ \sigma_N\, x_i^{(N)} \end{bmatrix} - \underline{\mu}\right)$$

as the final DM approximation of the wanted Gaussian density

$${}^{\mathrm{G}}f_N\Big(\underline{x}^{(1:N)}\Big) = \mathcal{N}\Big(\underline{x}^{(1:N)} - \underline{\mu},\ \mathbf{C}\Big) \quad .$$



(a) S^{2}KF (state-of-the-art) (b) Conditional (proposed) (c) Online (optimal but more complex)

Fig. 4: Visualization of three different Gaussian sampling schemes.
Yellow: standard normal distribution.
Red: one axis re-scaled to 10 %.
Blue: one axis completely collapsed.
Samples are shown as points, density contour lines enclosing 95 % of the probability mass are shown as dashed lines (circle or ellipse). Vertical grey lines show the paths along which the points move during rescaling in (a) and (b).
(a) Standard normal Gaussian sampling with subsequent re-scaling, as used in the S^{2}KF.
(b) Proposed conditional sampling of standard normal distribution with subsequent re-scaling along principal axes.
(c) Online sampling, i.e., numerically optimizing the distance measure separately for each of the three Gaussians.
Note that in (a), the standard normal samples (yellow) are optimal. In (b), the marginal distribution (blue) is optimal, and the scaled distribution (red) is suboptimal but better than in (a). In (c), all three DM distributions are optimal.

VI. Density Function Distance

We will now sketch the distance measure D_n that we use to compare DM densities to Gaussian densities. For a detailed introduction and motivation of the distance measure, see [27]. For more details on DM approximation of Gaussian densities using this distance measure, see [28] and [29].

Distance measures between continuous densities (like Gaussians) and discrete densities on continuous support (like DMs) are not straightforward to define. Typical distance measures such as the L_2 norm or the Kullback–Leibler divergence fail here, because it is not feasible to simply compare density function values as both functions do not share a common support. Instead, the local *point density* of samples (DM components) must be taken into account. Therefore, in a first step, a transformation (the LCD) using kernel functions is employed, ensuring common support. An L_2-based distance is then calculated between these transformed distributions.

A. Localized Cumulative Distribution (LCD)

First, each density function $f(\underline{x})$ is transformed to its LCD $F(\underline{m},\, b)$ that is defined as

$$F(\underline{m}, b) = \int_{\mathbb{R}^n} f(\underline{x}) \cdot K(\underline{x} - \underline{m},\, b)\ \mathrm{d}\underline{x}\ ,$$

with an isotropic Gaussian kernel $K(\underline{x} - \underline{m}, b)$

$$K(\underline{x} - \underline{m}, b) = \prod_{k=1}^{n} \exp\left\{ -\frac{1}{2} \frac{\left(x^{(k)} - m^{(k)}\right)^2}{b^2} \right\}\ .$$

Intuitively, this can be seen as a convolution of the density function with a Gaussian kernel with variable kernel width b.

B. A Modified Cramér-von Mises Distance

Second, a distance measure $D_n(\cdot,\cdot)$ between the LCD ${}^{\mathrm{G}}F_n^{\mathrm{s}}(\underline{m}, b)$ of the Gaussian density and the LCD ${}^{\mathrm{DM}}F_n^{\mathrm{s}}(\underline{m}, b)$ of the DM density is defined

$$D_n = \int_{\mathbb{R}^+} w(b) \cdot \int_{\mathbb{R}^n} \left({}^{\mathrm{G}}F_n^{\mathrm{s}}(\underline{m}, b) - {}^{\mathrm{DM}}F_n^{\mathrm{s}}(\underline{m}, b) \right)^2\ \mathrm{d}\underline{m}\, \mathrm{d}b\ , \tag{7}$$

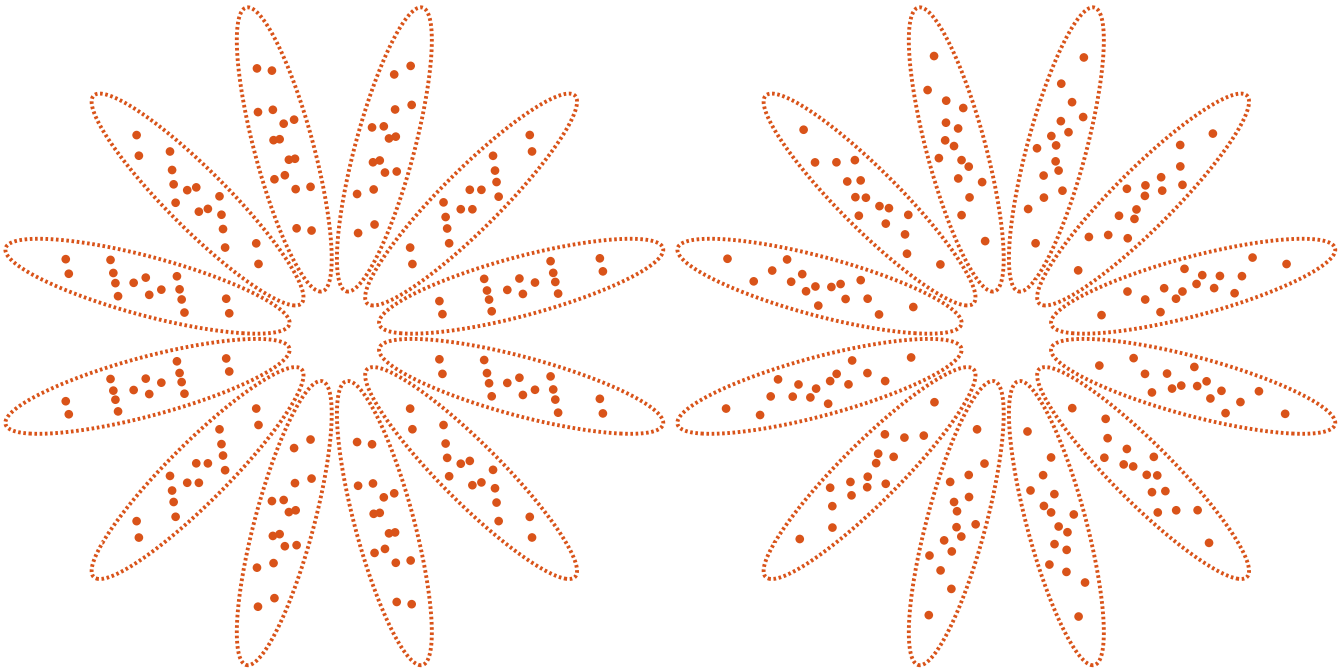
where $w(b)$ is a weighting function, here $w(b) = b^{-(n-1)}$.

The inner integral is the L_2 norm between both LCDs for a fixed kernel width b. The outer integral collects information regarding the distance from many different kernel widths into one scalar value, D_n.

Closed form solutions of the integrals exist for standard normal distributions [28].

C. Nonlinear Optimization

The distance measure (7) is the objective function of the nonlinear optimization problem in (1), (2), and (5) yielding



(a) S^2KF (state-of-the-art) (b) Conditional (proposed)

Fig. 5: Various Gaussian samplings with different rotational part $\mathbf{R}$ in the covariance matrix $\mathbf{C} = \mathbf{R}\boldsymbol{\Sigma}\mathbf{R}^\top$. Standard deviations along principal axes are 1 and 0.2 in each case, i.e., $\boldsymbol{\Sigma} = \mathrm{diag}(1, 0.04)$. Mean vectors $\underline{\mu}$ are chosen such that all Gaussians can be shown without overlapping. (a) For state-of-the-art S^2KF sampling, arrangement of points and sampling quality clearly depend on $\mathbf{R}$. (b) With the proposed conditional sampling, the achieved sampling quality is invariant against $\mathbf{R}$.

the desired deterministic sample locations. We used stochastic Gaussian samples as initial guess. For optimization, we applied the MATLAB function `fminunc()` that uses a BFGS Quasi-Newton method.

VII. Evaluation

In this section, we provide visual examples of various deterministic sampling schemes. This helps to compare the proposed deterministic conditional sampling to existing approaches.

A. Anisotropic Re-Scaling of Deterministic Samples

The effect of anisotropic re-scaling on deterministic standard normally distributed samples is visualized in Fig. 4a for the existing S^2KF approach and in Fig. 4b for the proposed conditional sampling scheme. Yellow points represent the standard normal distribution as stored in the sample library for 16 samples in two dimensions in this case. Covariance adaptation is then applied such that the standard deviation of one axis is reduced from 1 to 0.1, see red points. Blue points show the extreme case where the standard deviation of one axis is reduced to 0, as is the case for singular, positive semidefinite covariance matrices.

Clearly the conditional sampling scheme provides a more uniform arrangement of samples than the S^2KF approach while causing just about the same online computational load and even reduced offline computational load and library size.

If the LCD sampling is performed entirely online, separately for each specific Gaussian density, see Fig. 4c, the sampling quality is of course better than the proposed conditional sampling or S^2KF sampling. However, the online computational load for this scheme is much higher and not well suited for real-time applications.

B. Rotation of Deterministic Samples

In Sec. VII-A, we visualized the effect of anisotropic re-scaling encoded in the diagonal matrix $\boldsymbol{\Sigma}$. Yet the full covariance matrix $\mathbf{C}$ may also contain some rotational part $\mathbf{R}$ (6), which will be discussed here.

Fig. 5 shows various instances of deterministic sampling of Gaussian densities with identical eigenvalues $\boldsymbol{\Sigma}$ but different rotational parts $\mathbf{R}$. With S^2KF sampling in Fig. 5a, the samples are obviously arranged differently for different $\mathbf{R}$. The relevant state space that contains significant probability mass (inside the dashed ellipses) is not always covered uniformly. In the conditional sampling scheme however, $\boldsymbol{\Sigma}$ and $\mathbf{R}$ are applied separately and in a way that is tailored to the standard normal reference sample set. Therefore the arrangement of samples after covariance adaptation is invariant against $\mathbf{R}$ which is generally desirable. Overall distribution of samples, see Fig. 5b, appears to be relatively smooth and superior to S^2KF sampling.

VIII. Conclusions

We have proposed an efficient, real-time capable deterministic sampling scheme for multivariate Gaussian densities.

Compared to similar existing sampling schemes, it has provided superior coverage of the relevant state space. Additionally, and in contrast to the state-of-the-art S^2KF sampling scheme, the arrangement of samples is invariant against the rotational part, i.e., the eigenvectors of the covariance matrix. See Fig. 5 for a comparison.

Computational complexity for online sampling of arbitrary Gaussian densities has not increased significantly compared to S^2KF sampling, as we only need the eigenvalues and eigenvectors of the covariance matrices instead of their Cholesky factorization. Computational load and storage space for building the required offline library have even been greatly reduced.

We are currently including the proposed sampling scheme into the nonlinear estimation toolbox [32]. Instead of MATLAB's `fminunc()`, this library uses C++ code and MEX files to perform Quasi-Newton optimizations much faster.

References

[1] J. Steinbring, M. Pander, and U. D. Hanebeck, "The Smart Sampling Kalman Filter with Symmetric Samples," *Journal of Advances in Information Fusion*, vol. 11, no. 1, pp. 71–90, Jun. 2016.

[2] J. Steinbring, A. Zea, and U. D. Hanebeck, "Semi-Analytic Progressive Gaussian Filtering," in *Proceedings of the 2016 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2016)*, Baden-Baden, Germany, Sep. 2016.

[3] J. Steinbring and U. D. Hanebeck, "GPU-Accelerated Progressive Gaussian Filtering with Applications to Extended Object Tracking," in *Proceedings of the 18th International Conference on Information Fusion (Fusion 2015)*, Washington D.C., USA, Jul. 2015.

[4] ——, "Progressive Gaussian Filtering Using Explicit Likelihoods," in *Proceedings of the 17th International Conference on Information Fusion (Fusion 2014)*, Salamanca, Spain, Jul. 2014.

[5] U. D. Hanebeck, "PGF 42: Progressive Gaussian Filtering with a Twist," in *Proceedings of the 16th International Conference on Information Fusion (Fusion 2013)*, Istanbul, Turkey, Jul. 2013.

[6] M. F. Huber and U. D. Hanebeck, "Gaussian Filter based on Deterministic Sampling for High Quality Nonlinear Estimation," in *Proceedings of the 17th IFAC World Congress (IFAC 2008)*, vol. 17, no. 2, Seoul, Republic of Korea, Jul. 2008.

[7] J. H. Kotecha and P. M. Djuric, "Gaussian Sum Particle Filtering," *IEEE Transactions on Signal Processing*, vol. 51, no. 10, pp. 2602–2612, Oct. 2003.

[8] D. Frisch and U. D. Hanebeck, "Progressive Bayesian Update Using Interleaved Gaussian Mixture and Dirac Mixture," in *Proceedings of the 23rd International Conference on Information Fusion (Fusion 2020)*, Virtual, Jul. 2020.

[9] C. Stachniss, G. Grisetti, W. Burgard, and N. Roy, "Analyzing Gaussian Proposal Distributions for Mapping with Rao-Blackwellized Particle Filters," in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct. 2007, pp. 3485–3490.

[10] P. Bunch and S. Godsill, "Particle filtering with progressive Gaussian approximations to the optimal importance density," in *2013 5th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Dec. 2013, pp. 360–363.

[11] U. D. Hanebeck and M. Pander, "Progressive Bayesian Estimation with Deterministic Particles," in *Proceedings of the 19th International Conference on Information Fusion (Fusion 2016)*, Heidelberg, Germany, Jul. 2016.

[12] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, Jul. 2000. [Online]. Available: https://doi.org/10.1023/A:1008935410038

[13] C. Chlebek, A. Hekler, and U. D. Hanebeck, "Stochastic Nonlinear Model Predictive Control Based on Progressive Density Simplification," in *Proceedings of the 51st IEEE Conference on Decision and Control (CDC 2012)*, Maui, Hawaii, USA, Dec. 2012.

[14] M. Dolgov, G. Kurz, D. Grimm, F. Rosenthal, and U. D. Hanebeck, "Stochastic Optimal Control Using Local Sample-Based Value Function Approximation," in *Proceedings of the 2018 American Control Conference (ACC 2018)*, Milwaukee, Wisconsin, USA, Jun. 2018.

[15] D. J. Higham., "An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations," *SIAM Review*, vol. 43, no. 3, pp. 525–546, Aug. 2001. [Online]. Available: https://doi.org/10.1137/S0036144500378302

[16] S. Au and J. Beck, "Important sampling in high dimensions," *Structural Safety*, vol. 25, no. 2, pp. 139 – 163, Apr. 2003. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167473002000474

[17] E. Bertschinger, "Multiscale Gaussian Random Fields and Their Application to Cosmological Simulations," *The Astrophysical Journal Supplement Series*, vol. 137, no. 1, pp. 1–20, Nov. 2001. [Online]. Available: https://doi.org/10.1086%2F322526

[18] H. Rue, "Fast sampling of Gaussian Markov random fields," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 63, no. 2, pp. 325–338, 2001. [Online]. Available: https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/1467-9868.00288

[19] K. Li, G. Kurz, L. Bernreiter, and U. D. Hanebeck, "Simultaneous Localization and Mapping Using a Novel Dual Quaternion Particle Filter," in *Proceedings of the 21st International Conference on Information Fusion (Fusion 2018)*, Cambridge, United Kingdom, Jul. 2018.

[20] K. Li, F. Pfaff, and U. D. Hanebeck, "Geometry-Driven Stochastic Modeling of SE(3) States Based on Dual Quaternion Representation," in *Proceedings of the 2019 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2019)*, Taipei, Republic of China, May 2019.

[21] K. Li, D. Frisch, S. Radtke, B. Noack, and U. D. Hanebeck, "Wavefront Orientation Estimation Based on Progressive Bingham Filtering," in *Proceedings of the IEEE ISIF Workshop on Sensor Data Fusion: Trends, Solutions, Applications (SDF 2018)*, Oct. 2018.

[22] K. Li, D. Frisch, B. Noack, and U. D. Hanebeck, "Geometry-Driven Deterministic Sampling for Nonlinear Bingham Filtering," in *Proceedings of the 2019 European Control Conference (ECC 2019)*, Naples, Italy, Jun. 2019.

[23] K. Li, F. Pfaff, and U. D. Hanebeck, "Hyperspherical Deterministic Sampling Based on Riemannian Geometry for Improved Nonlinear Bingham Filtering," in *Proceedings of the 22nd International Conference on Information Fusion (Fusion 2019)*, Ottawa, Canada, Jul. 2019.

[24] S. J. Julier and J. K. Uhlmann, "New extension of the Kalman filter to nonlinear systems," in *Signal processing, sensor fusion, and target recognition VI*, vol. 3068. International Society for Optics and Photonics, Jul. 1997, pp. 182–193.

[25] S. J. Julier, "The scaled unscented transformation," in *Proceedings of the 2002 American Control Conference (IEEE Cat. No.CH37301)*, vol. 6, May 2002, pp. 4555–4559 vol.6.

[26] U. D. Hanebeck, "Deterministic Sampling of Multivariate Densities based on Projected Cumulative Distributions," in *Proceedings of the 54th Annual Conference on Information Sciences and Systems (CISS 2020)*, Princeton, New Jersey, USA, Mar. 2020.

[27] U. D. Hanebeck and V. Klumpp, "Localized Cumulative Distributions and a Multivariate Generalization of the Cramér-von Mises Distance," in *Proceedings of the 2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2008)*, Seoul, Republic of Korea, Aug. 2008, pp. 33–39.

[28] U. D. Hanebeck, M. F. Huber, and V. Klumpp, "Dirac Mixture Approximation of Multivariate Gaussian Densities," in *Proceedings of the 2009 IEEE Conference on Decision and Control (CDC 2009)*, Shanghai, China, Dec. 2009.

[29] I. Gilitschenski and U. D. Hanebeck, "Efficient Deterministic Dirac Mixture Approximation," in *Proceedings of the 2013 American Control Conference (ACC 2013)*, Washington D.C., USA, Jun. 2013.

[30] J. Steinbring and U. D. Hanebeck, "LRKF Revisited: The Smart Sampling Kalman Filter (S2KF)," *Journal of Advances in Information Fusion*, vol. 9, no. 2, pp. 106–123, Dec. 2014. [Online]. Available: {http://confcats_isif.s3.amazonaws.com/web-files/journals/entries/441_1_art_11_17020.pdf}

[31] ——, "S2KF: The Smart Sampling Kalman Filter," in *Proceedings of the 16th International Conference on Information Fusion (Fusion 2013)*, Istanbul, Turkey, Jul. 2013.

[32] J. Steinbring, "Nonlinear Estimation Toolbox." [Online]. Available: https://bitbucket.org/nonlinearestimation/toolbox